

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 5, 2003, 01:50:12 ; Search time 287 seconds  
(without alignments)  
3452.541 Million cell updates/sec

Title: US-09-847-081b-2

Perfect score: 2270

Sequence: 1 MSMSVALLWVSPTEVSENG.....IYAKSLVPPNRTSSPLAKT 440

Scoring table: BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Xgapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cgn2\_1/USPTO.spool/US09847081/runat\_01042003.120128.26729/app\_query.fasta\_1.583  
-DB-N Geneseq 101002 -OFMT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=100 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptn -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09847081.ecgn\_1.1.396.runat\_01042003.120128.26729 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.:

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	2270	100.0	1728	24	AAI66366
2	1826	80.4	1826	19	AAV03880
3	1801	79.3	1712	24	AAI66367
4	1799	79.3	1814	19	AAV03881
5	1774	78.1	1795	19	AAV03878
6	1762	77.6	1316	19	AAV03879
7	1751	77.1	1591	17	AAQ99323
8	1744	76.8	1239	19	AAV17247
9	1737	76.5	1239	21	AAZ99482
10	1729	76.2	1646	12	AAQ12495
11	1595	70.3	2868	19	AAV16951
12	1577	68.6	1921	19	AAV16949
13	1542	67.9	1566	21	AAAC48162
14	1542	67.9	1703	21	AAAC35120
15	1531	67.4	2085	19	AAV16948
16	1523	67.1	1932	19	AAV16950
17	1514	66.7	2585	23	ABLA1600
18	1431.5	63.1	1304	24	ABA97361
19	1393	61.4	1397	21	AAZ29145
20	1252.5	55.2	1448	21	AAZ29139
21	1078	47.5	1021	21	AAZ29146
22	1033.5	45.5	3485	13	AAQ29121
23	1033	45.5	992	21	AAZ29144
24	992	43.7	1060	21	AAZ29143
25	781.5	34.4	888	21	AAZ29140
26	511	22.5	476	21	AAZ29142
27	468	20.6	684	22	AAH44248
28	458.5	20.2	1509	15	AAQ64910
29	442	19.5	749	19	AAV03882
30	420	18.5	766	21	AAZ29141
31	378	16.7	5609	21	AAAB1459
32	378	16.7	349980	21	AAF21609
33	378	16.7	1437668	21	AAA81490
34	364	16.0	1198	12	AAQ13718
35	364	16.0	1198	17	AAAT40791
36	364	16.0	1198	17	AAAT41743
37	364	16.0	1198	17	AAAT37093
38	364	16.0	1198	18	AAAT91543
39	363.5	16.0	930	18	AAAT66534
40	350.5	15.4	1232	19	AAV19122
41	350.5	15.4	1232	20	AAZ25063
42	350.5	15.4	1232	24	ABA97354
43	349.5	15.4	6918	11	AAQ06299
44	347.5	15.3	1386	13	AAQ29122
45	344.5	15.2	930	18	AAZ23885

ALIGNMENTS

RESULT 1	
AAI66366	
ID	AAI66366 standard; cDNA; 1728 BP.
XX	
AC	AAI66366;
XX	
DT	29-JAN-2002 (first entry)
XX	
DE	Nicotiana tabacum phytoene synthase coding sequence #1.
XX	
KW	Phytoene synthase; zeta carotene desaturase; herbicide; transgenic plant;
KW	plant growth regulator; herbicidal; tobacco; ss.
XX	
OS	Nicotiana tabacum.
XX	
Key	Location/Qualifiers
FT	244...1566
FT	/*tag- a

/product= "phytoene synthase"

FT XX DE10022362-A1.  
 PN XX 15-NOV-2001.  
 PD XX  
 XX XX 08-MAY-2000; 2000DE-1022362.  
 PF XX  
 XX XX 08-MAY-2000; 2000DE-1022362.  
 PR XX  
 XX XX (FARB ) BAYER AG.  
 PA XX  
 PI XX Busch M, Hain R;  
 XX XX  
 DR XX WPI: 2002-027336/04.  
 DR XX P-PSDB; AM51841.  
 XX XX

New nucleic acid encoding tobacco zeta-carotene desaturase, useful for screening compounds with herbicidal activity -  
 XX  
 PS Claim 14; Page 12-17; 44pp; German.  
 XX

The present invention provides the protein and coding sequences of phytoene synthases and zeta-carotene desaturase from *Nicotiana tabacum*. The sequences can be used to identify compounds capable of altering the expression of these genes, which are therefore useful as plant growth regulators and herbicides. They can also be used to produce transgenic plants. The present sequence is the coding sequence of a tobacco phytoene synthase.  
 CC  
 XX

Sequence 1728 BP; 513 A; 290 C; 433 G; 492 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 7,898-234 Length: 1728  
 Score: 2270.00 Matches: 440  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-847-081B-2 (1-440) x AAI66366 (1-1728)

QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20  
 DB 244 ATGAGCATGCTGCTTGTGTGTGGTGTGTTCTCCACATCCGAGGCTCGAATGG 303  
 QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40  
 DB 304 ACAGGATTGTGGATTCAGTCCGAGAAAGAAACCCGCTCTTTGTATCATCCAGGTTCTTA 363  
 QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuLysGlyGlyArgGlnArgTrp 60  
 DB 364 GCTCGAGATAGGAATTTGATGTGGATGGAGATCAAGAAAGGTGGGAGACAAAGGTGG 423  
 QY 61 AsnPhcGlySerLeuLeuAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80  
 DB 424 AATTTTGGCTCTTAACTGCTGATCCAGATATCATGCTTGGTGGATCAAGAAGTCAA 483  
 QY 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100  
 DB 484 AAGGAAGACATTTCTCTGATAGTCCAGTCTGGTGGCTAGCCAGCTGGAGAAATGACT 543  
 QY 101 ValSerSerGluLysLysValTrpAspValValLeuLysGlnAlaAlaLeuValLysArg 120  
 DB 544 GTGTCATCAGAGAAAGGTGATGATGATGATGATTAAGCAGGCGCTTTAGTAGAGAGG 603  
 QY 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeu 140  
 DB 604 CAGCTCAGATCTACCCATGATTTAGAAAGTGAAGCCGATATTGTTTCCAGGGAATTG 663  
 QY 141 GlyLeuLeuSerCysAlaLysArgCysGlyGluValCysAlaGluTrpAlaLysThr 160  
 DB 664 GGCTTTTGGTGAAGCATATGATCGTTGTGGGAAGTATGTGCAGAGTATGCAAGAGACA 723

161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyr 180  
 DB 724 TTTTACTTAGGAACCAAGCTAATGACCCAGAGAGAGAGAGCTATCTGGGCAATATAT 783  
 QY 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200  
 DB 784 GTGTGGTCCAGGAGACGATGAGCTTGTGTGGCCCTTAATGCATCCACATAACTCCG 843  
 QY 201 GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220  
 DB 844 CAAGCTTTAGATAGGTGGGAGACAGGCTGGAGATATTTTCAGTGGCGGCCATTGAT 903  
 QY 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240  
 DB 904 ATGCTTGATGCTGCTTTATCCGATACTGCTCCAGATTTCTGTTGATATTTCAGCATTC 963  
 QY 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAsp 260  
 DB 964 AGAGATATGATTGAAGGAATGGTATGGACTTGTGGAAATCCAGATACAAAACCTTCGAT 1023  
 QY 261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280  
 DB 1024 GAGCTATATCTCTATTGTTACTATGTTGCTGCTAGTGTAGGATTTGATGCTTCCAGTT 1083  
 QY 281 MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAla 300  
 DB 1084 ATGGGTATTGCACTGGAATCAAGGCAACACAGAGATGTATATATGCTGCTTTGGCT 1143  
 QY 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320  
 DB 1144 TTAGGCTTGCATCAATCACTAATCAATATATCTCAGAGATGTAGGAGAAAGATGCCAGAGA 1203  
 QY 321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340  
 DB 1204 GGAAGAGTATATCTGCTCAAGATGAATTAGCACAGGCGGCTCTCCGACGAGACATA 1263  
 QY 341 PheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAla 360  
 DB 1264 TTTGCTGGAAAGATGACTATAGTGGAGAACTTTATGAAGAAACAAATTCAGAGGGCG 1323  
 QY 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380  
 DB 1324 AGAAATCTTTGATGAGTCAGAGAAAGGTGTACAGAACTGGACCTCTGCTAGTAGTGG 1383  
 QY 381 ProValLeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp 400  
 DB 1384 CCTGTGTTAACAGCGCTGCTGTGTATCGCAAGATATTGGACGAGATTGAAGCCACGAC 1443  
 QY 401 TyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPro 420  
 DB 1444 TACACAACTTCACAGGAGGCTTATGTAGCAAGCCAAAGAGCTTCTCACCCTTGCCC 1503  
 QY 421 IleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLysThr 440  
 DB 1504 ATTGCTTATGCAAAATCTCTTGTGCCCTTAATAGAATCTCTCTCCACTAGCAAGACA 1563

RESULT 2  
 AAV03880  
 ID AAV03880 standard; cDNA; 1826 BP.  
 XX  
 AC AAV03880;  
 XX  
 XX 29-APR-1998 (first entry)  
 XX  
 DE Phytoene synthase coding sequence from *N. tabacum*.  
 XX  
 KW Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;  
 KW ultra violet absorber; food colour; ss.  
 XX  
 OS *Nicotiana tabacum*.  
 XX  
 FH Key Location/Qualifiers



KW Phytoene synthase; zeta: carotene desaturase; herbicide; transgenic plant;  
 KW plant growth regulator; herbicidal; tobacco; ss.  
 OS Nicotiana tabacum.

XX Key Location/Qualifiers  
 XX CDS 333..1565  
 XX FT /\*tag-a  
 XX FT /product- "phytoene synthase"

XX DE10022362-A1.

XX PN 15-NOV-2001.

XX PD 08-MAY-2000; 2000DE-1022362.

XX PF 08-MAY-2000; 2000DE-1022362.

XX PR (FARB ) BAYER AG.

XX PA Busch M, Hain R;

XX PI WPI; 2002-027336/04.

XX DR P-PSDB; AAM51842.

XX PT New nucleic acid encoding tobacco zeta-carotene desaturase, useful for  
 XX screening compounds with herbicidal activity -

XX PS Claim 14; Page 21-26; 44pp; German.

XX The present invention provides the protein and coding sequences of  
 CC phytoene synthases and zeta-carotene desaturase from Nicotiana tabacum.  
 CC The sequences can be used to identify compounds capable of altering the  
 CC expression of these genes, which are therefore useful as plant growth  
 CC regulators and herbicides. They can also be used to produce transgenic  
 CC plants. The present sequence is the coding sequence of a tobacco  
 CC phytoene synthase.

XX SQ Sequence 1712 BP; 540 A; 265 C; 421 G; 475 T; 11 other;

#### Alignment Scores:

Pred. No.: 2,04e-183 Length: 1712  
 Score: 1801.00 Matches: 357  
 Percent Similarity: 87.93% Conservative: 29  
 Best Local Similarity: 81.34% Mismatches: 24  
 Query Match: 79.34% Indels: 29  
 DB: 24 Gaps: 4

US-09-847-081B-2 (1-440) x AAI66367 (1-1712)

QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22  
 Db ATGCTGTTGCTTGTATGGTGTGTTTACCT--TGTAAGTCTCAATGGGACAGA 389  
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42  
 Db TTCTTGATTCAGTCGCGGAGGACCGGTTTGTATGCTCG-----AGG 437  
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgLysLysGlyArgGlnArgTrpAsnPhe 62  
 Db CATAGGAATTTAGTGTGCAATGAGAAACAAAGAGAGGTGTGAACAAAGGTGTAATTT 497  
 QY 63 GlySerLeuIleAlaAppProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82  
 Db GGT----- 500  
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101  
 Db TCTAGAGTCTGCTATGGTGGTACACCGCGGGGAGAAATGGGACGATG 551  
 QY 102 SerSerGluLysLysValTrpAspValValLeuLysGlnAlaLeuValLysArgGln 121  
 Db ACATCAGACAGATGGTGTATGATGGTGTGTTTAAACAAAGCAGCTTTAGTGAAGGCGAG 611

QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGly 141  
 Db TTGAGATCTGCTGATGATTTAGAGTGAAGCGGAGATCCTCTCCCGGAATTTGAGC 671  
 QY 142 LeuLeuSerGluAlaTrpAspArgCysGlyGluValCysAlaGluTrpAlaLysThrPhe 161  
 Db TTGTTGAGTGAAGCATATGATAGGTAGTAAAGTATGTCAGAGTATGCAAGACATTT 731  
 QY 162 TyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrVal 181  
 Db TACTTHGGAAACCATGYTAATGACTCCAGAGAGAAAGGCGCTATTGGGCAATATATGTG 791  
 QY 182 TrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201  
 Db TGGTCAGGAGAACAGATGAACCTTGTATGGCCAAACGATCATATATACACCCCAA 851  
 QY 202 AlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221  
 Db GCCTTAGATAGGTGGGAGACCGGCTTGAAGATGTTTTCAGCGGGGCGACCATTTGATATG 911  
 QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241  
 Db CTCGATGCTGCTTGTCCGATAGTCTTCCAACTTCCAGTTGATATTCAGCCGTTTCTG 971  
 QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGlu 261  
 Db GATATGATTGAAGGAATGCGTATGAGTGTGAGGAAGTCAAGATATAGAACTTTGATGAG 1031  
 QY 262 LeuTyrLeuTyrCysTyrTrpValAlaGlyThrValGlyLeuMetSerValProValMet 281  
 Db CTTTACCTCTATGTTTATTACGTTGCTGTTGTTGGTGTGATGAGTGTTCCTCAATATG 1091  
 QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTrpAsnAlaAlaLeuAlaLeu 301  
 Db GGTATTGACCTGATTCAAAGGCAACACAGACAGCGGTATATATGACGCTTTGGCTTA 1151  
 QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321  
 Db GGAATCGCAATCAACTAACGACATCTACAGAGATGTTGGGAAGATGCCAGAAGAGGA 1211  
 QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341  
 Db ACAGTCTACTTACCTCAAGATGAATTTAGCACAGCGAGGCTCTCTCGACGATGACATATT 1271  
 QY 342 AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361  
 Db GCTGGAAGAGTGACTGATAAGTGGAAGAGCTTTATGAAGAAGCAAAATCCAGAGGCAAGA 1331  
 QY 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381  
 Db AAGTTCTTCGATGAGGACAGGAGGAGGTACACAACTAGCTAGCTCAGCTAGCATGCCT 1391  
 QY 382 ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyr 401  
 Db GTATGGGCATCTTGTGTTTACCCCAATACTGACGAGATTGAAGCAATGACTAC 1451  
 QY 402 AsnAsnPheThrArgArgAlaTrpValSerLysProLysLysLeuLeuThrLeuProIle 421  
 Db AACAACTTCACAAAGAGAGCTTATGAGCAACCAACCAAGAGCTAATTTCTTACCTATT 1511  
 QY 422 AlaTyrAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys 439  
 Db GCTTTATGCAAAATCTCTTGTGCCCCCTACAAAGAAGCTTTGTCTACCTCTAGCTAAG 1566

#### RESULT 4

AAV03881  
 ID AAV03881 standard; cDNA; 1814 BP.  
 XX  
 AC AAV03881;  
 XX  
 DT 29-APR-1998 (first entry)  
 XX













Db	589	GACGGAGCACTTAGTGACACTGTGAGCAATTCCTCGTGGACATCCAAACCTTTTCGGGAC	648
Qy	243	MetileGlUglyeArgetMetAspLeuThrLysSerArgTyrlYstThrPheAspGluleu	262
Db	649	ATCATCGAGGCATGAGAATGATCTTCGTAACTCCTGTAAGAATTITGATGAGTTG	708
Qy	263	TyrLeuTyrcystTyrrValalaglyThrValGlyLeuMetSerValProValMetGly	282
Db	709	TATTTGCTACTGCTACTACGTGCGAGAACCGGTGGGCCCTTATGTCACTGCCTCATCGGA	768
Qy	283	IleAlaproGluSerLysAlathrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly	302
Db	769	ATTGACCAGAGNTAAGCTACTAGTACTGTTTACCCGACGACATAGCAATTAGGT	828
Qy	303	LeuAlaasnGlnLeuThrasnIleLeuArgaspValGlyGluaspAlaArgArgGlyArg	322
Db	829	ATACTTAACACAGCTTACAATAATCTTTGAGGACGTGGGTGAGGACGACGTAGGGGTCTG	888
Qy	323	ValtyrLeuproGlnaspgluleualGlnalaglyLeuSeraspGluaspIlePheAla	342
Db	889	GTGTATCTCCACAGACGAGCTCGCTCAAGCTGGATTGAGTGACGAGGACATTTTCGCA	948
Qy	343	GlyArgvalThrAsplystrpargasnPhemetLysLysGlnIleGlnArgAlaArgLys	362
Db	949	GGTCGTGTTACAGACAAGTGGAGGATTTTCATGAAAAAGCAGATTCACCGTGCCTGAAA	1008
Qy	363	PhePheaspGluSerClulysglyValthrGluLeuaspsSerAlaSerArgTrpProVal	382
Db	1009	TTTTTCGAGAGAGCTGAAAGGGAGTTACTGAGCTTCTTAGTGATCAAGGTTTCCAATT	1068
Qy	383	LeuthrAlaleuLeuThryrArgLyslleLeuaspgluilleGluAlaAsnAspTyrAsn	402
Db	1069	TGGGCCAGCCTTGCTCTATAGAAAGATTTTGGACGAAATCGAGGCTAACGATTATAAT	1128
Qy	403	AsnPheThrArgArgAlatyrrValSerLysProLysLysLeuLeuThrLeuProIleAla	422
Db	1129	AATTTTACTAAAGCTGCTTACGTTTCTAAGACGAAAAAATATTCGCTCTTCCAATGCT	1188
Qy	423	TyrAlalySerLeuValProProasnargThrserSer 435	
Db	1189	TACGCTAAGAGCTTGGTCCACCACTAAGACAGCTAGC 1227	
RESULT 9			
AAZ99482			
ID	RAZ99482	standard; cDNA; 1239 BP.	
XX	AC	AAZ99482;	
XX	DT	03-JUL-2000 (first entry)	
XX	DE	cDNA encoding a phytoene synthase polypeptide.	
KW	KW	Gibberellic acid; copalyl diphosphate synthase; 3beta-hydroxylase;	
KW	KW	2-oxidase; phytoene synthase; C-20 oxidase; 2beta,3beta-hydroxylase;	
KW	KW	seed germination; seedling growth; gibberellin biosynthetic pathway;	
XX	XX	transgenic plant; hypocotyl; epicotyl; ss.	
OS	OS	Lycopersicon esculentum.	
XX	FX	Key Location/Qualifiers	
FT	FT	1..1239	
FT	FT	/tag= a	
FT	FT	/product= "phytoene synthase"	
FT	FT	/transl_except= (pos: 1027..1029, aa: Lys)	
PN	PN	/transl_except= (pos: 1057..1059, aa: Arg)	
XX	XX	WO200009722-A2.	
PD	PD	24-FEB-2000.	
XX	XX	10-AUG-1999; 99WO-US18066.	
XX	XX		



```

Db 543 GCGCTGTGAGTGAAGCATATGATAGGTGGTGAAGTATGTCGAGATATGCAAGACG 602
Qy 161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTyrAlaIleTyr 180
Db 603 TTTAACTTAGCAACTATGCTAATGACTCCCGAGAGAAGGGCTATCTGGTCAATATAT 662
Qy 181 ValTyrCysArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200
Db 663 GTATGGTGCAGAGAGACAGATGAACCTGTTGATGGCCCAACGCATCATATATACCCG 722
Qy 201 GlnAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSerGlyAlaGProPheAsp 220
Db 723 GCAGCCTTAGATAGTGGGAAATAGGCTAGAGATGTTTCAATGGCGGCCATTTGAC 782
Qy 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240
Db 783 ATGCTCGATGGTCTTGTGTCGATACAGATTTCTAACTTCCAGTTGATATTCAGCCATTC 842
Qy 241 ArgAspMetIleGluGlyMetArgMetAspLeuTyrLysSerArgTyrLysThrPheAsp 260
Db 843 AGCATATGATGANGAATGCGTATGGACTTGAGAAAATCGAGATACAAAACCTTCGAC 902
Qy 261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
Db 903 GAACATACCTTTATGTTATGTTGCTGGTACGTTGGGTGATGAGTGTTCCAATT 962
Qy 281 MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAla 300
Db 963 ATGGGTATCGCCCTGAATCAAGGCAACACAGAGACGCTATATATGCTGCTTGGCT 1022
Qy 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArg 320
Db 1023 CTGGGGATCGCAATCAATTAACCTACATCTCAGAGATGTTGGAGAGATGCCAGAGA 1082
Qy 321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340
Db 1083 GGAAGAGTCTACTTGCTCCTCAAGATGAATTAGCACAGCAGCTCTATCCGATGAATATA 1142
Qy 341 PheAlaGlyArgValThrAspLysTyrArgAsnPheMetLysLysGlnIleGlnArgAla 360
Db 1143 TTTCTGGAAGGGTACCGCATAAATGAGATCTTTATGAAGAAACAATACATAGGCA 1202
Qy 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTyr 380
Db 1203 AGAAGTCTTGTGATGAGGAGAGAGAGGCGTGACAGAAATGAGCTCAGCTAGTAGATTC 1262
Qy 381 ProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp 400
Db 1263 CCTGTATGGGCATCTTTGGTCTTGTACCGCAAAATACATAGATGAGATTGAAGCAATGAC 1322
Qy 401 TyrAsnAsnPheThrArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPr 420
Db 1323 TACAACAACTTCACAAAGAGAGCATATGTGACGAAATCAAGCAAGTGTATGTCATTACC 1382
Qy 420 oileAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLys 439
Db 1383 TATTGCATATGCAAAATCTCTGTGGCTCTCT- - - - -ACAAACTGCTCTCTTCAAA 1434

```

## RESULT 11

AAV16951

ID AAV16951 standard; cDNA to mRNA; 2868 BP.

XX AC

AAV16951;

XX DT

06-JUL-1998 (first entry)

XX DE

Nucleic acid encoding phytoene synthase 4.

XX KW

Phytoene synthase; breeding; variable flower colour; ds.

XX OS

Gentiana lutea.

XX

```

FH Key Location/Qualifiers
FT CDS 400..1689
XX /*tag= a
XX JP10084966-A.
XX 07-APR-1998.
XX 17-SEP-1996; 96JP-0245107.
XX 17-SEP-1996; 96JP-0245107.
XX (IWAT-) IWATE KEN.
XX WPI; 1998-264853/24.
XX P-PSDB; AAW46964.
XX Phytoene synthase gene - useful for breeding plant of variable
XX flower colour
XX Claim 4; Pages 12-14; 15pp; Japanese.
XX The present sequence encodes phytoene synthase 4. It was isolated from
XX a cDNA library prepared from mRNA extracted from the petals of Gentiana
XX lutea. The nucleic acid sequence was amplified from the library using
XX PCR primers AAV16952-53. The phytoene synthase gene is useful for
XX breeding plants with variable flower colours.
XX SQ Sequence 2868 BP; 887 A; 499 C; 648 G; 833 T; 1 other;

Alignment Scores:
Pred. No.: 6,03e-161 Length: 2868
Score: 1595.00 Matches: 313
Percent Similarity: 83.68% Conservative: 46
Best Local Similarity: 72.96% Mismatches: 60
Query Match: 70.26% Indels: 10
DB: 19 Gaps: 3
US-09-847-081B-2 (1-440) x AAV16951 (1-2868)
Qy 1 MetSerMetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGly 20
Db 394 GTTAACATGCTATTCTACGCTATGGTGTCTTCGCCGAGTCTCGAAGTTTGGAGTGC 453
Qy 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 454 AATGTTTTCTTGGAGCCAATTCGAGAAAGT-----TACCAT 489
Qy 41 AlaArgAspArgAsnLeuMetTyrAsnGlyArgIleLysLysGlyGlyArgGlnArgTyr 60
Db 490 TTTTCGGATAAAGCTTTAATGTACATGGAAGAGTTAAGAAAAGTAGACACCAAGGCGT 549
Qy 61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
Db 550 AGATCACGTTATGGGTGGAGATTGAGTTCAITTTGCTTGAGAGACTCTCGATTAGAG 609
Qy 81 ---LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet 99
Db 610 ACCCGGGAAGAAGATTATCGTATCTCCAGTATTATAGCTACCCCGCAGGAGAAATG 669
Qy 100 ThrValSerSerGluLysLysValTyrAspValValLeuLysGlnAlaLeuValLys 119
Db 670 ACGATGACATCAGAGCAAAAGGTTTATGATCGTCTTTAAAGCAAGCAGCTTTGATTAT 729
Qy 120 ArgGlnLeuArgSerThrAspAspLeuValLysProAspIleValValProGlyAsn 139
Db 730 AGACAGTGGAGTCTAGAGAAAATTTGGAGGTGAACCGGACATATTATTTGCCAGCAAC 789
Qy 140 LeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGlyTyrAlaLys 159
Db 790 GCGAACGCTGTTGAATGAAGCTTATGATCGGTGTCGAGAAGTATGTGCTGAATATCCCAAG 849
Qy 160 ThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleThrAlaIle 179

```

```

Db 850 TCATTCTACTGGGACCACCGCTCATGACACCGAGAGCGCTTACGTATCTGGCGGATA 909
Qy 180 TyrValTrpCysArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThr 199
Db 910 TATGATGCTGTAGGAGGACAGATGAGCTTGTGATGGCCTAACGGCTCACATRAAT 969
Qy 200 ProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPhe 219
Db 970 CCAACCGCGTGTAGATAGTGGGAAACCAAGATTAGAAGATGTTTCAAGGGCAACCTTT 1029
Qy 220 AspMetLeuAspAlaLeuSerAspThrValSerArgPheProValAspIleGlnPro 239
Db 1030 GATATGCTGTAGTGGCTTATCTGATACCAATACCAAGTATCTGTGGACATCCAGCCA 1089
Qy 240 PheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTrpLysThrPhe 259
Db 1090 TTTAGATATGATAGAGGAATCGGATGATCTGAGAAATCGAGATACAGAAATTC 1149
Qy 260 AspGluLeuTrpLysCysTrpTrpValAlaGlyThrValGlyLeuMetSerValPro 279
Db 1150 GATGAGCTGTATCTTACTGCTATATGCTGCTGCTACAGTTGGCTTGCATGAGTGACCA 1209
Qy 280 ValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTrpAsnAlaLeu 299
Db 1210 GTAATGGGCATGACCTGCACTAAGGCAACAACAGAAAGTGTATATGACAGCTTTA 1269
Qy 300 AlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArg 319
Db 1270 TCTTTGGGATCGGACACGCTGACTAACATCTTAAGGGATGTTGGAGAGATGCAAGA 1329
Qy 320 ArgGlyArgValTrpLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAsp 339
Db 1330 AGAGGAGAGTGTACCTACCTCAAGATGAATAGCACAGCAGGTTTATCAGATGAGGAC 1389
Qy 340 IlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArg 359
Db 1390 ATTTTGTCTGGAAGTGTACAGACAAATGGAGGATTTTATGAAGAACCAATCAAAAGG 1449
Qy 360 AlaArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArg 379
Db 1450 GCTAGAAATCTATGATGATGCAGAGAAAGAGTCCCGCAACTCAGCTCCGCGAGAGCAG 1509
Qy 380 TrpPro---ValLeuThrAlaLeuLeuTrpArgLysIleLeuAspGluIleGluAla 398
Db 1510 ATTGCTCTGTGGCGCAGCGTTGCTTTTATAGAAAAATATTGGATGAGTAGAAGCA 1569
Qy 399 AsnAspTrpAsnAsnPheThrArgArgAlaTrpValSerLysProLysLysLeuThr 418
Db 1570 AATGACTACACAAATTTACAAAGAGGCTTATGTAAACAGGCAAGAAAGCTATTAGCT 1629
Qy 419 LeuProIleAlaTrpAlaLysSerLeu 427
Db 1630 ATGCGCTGATGATGTCGCAAGTCTCTC 1656

```

## RESULT 12

AAV16949

ID AAV16949 standard; cDNA to mRNA; 1921 BP.

XX AC AAV16949;

XX DT 06-JUL-1998 (first entry)

DE Nucleic acid encoding phytoene synthase 2.

XX Phytoene synthase; breeding; variable flower colour; ds.

OS Gentiana lutea.

XX Key Location/Qualifiers

FH 412..1689

FT CDS

XX /\*tag= a

```

PN JP10084966-A.
XX 07-APR-1998.
XX 17-SEP-1996; 96JP-0245107.
XX 17-SEP-1996; 96JP-0245107.
XX (IWAT-) IWATE KEN.
XX WPI: 1998-264853/24.
XX P-PSDB; AAW46962.
XX Phytoene synthase gene - useful for breeding plant of variable
PT flower colour
XX Claim 2; Pages 7-9; 15pp; Japanese.
XX The present sequence encodes phytoene synthase 2. It was isolated from
CC a cDNA library prepared from mRNA extracted from the petals of Gentiana
CC lutea. The nucleic acid sequence was amplified from the library using
CC PCR primers AAV16952-53. The phytoene synthase gene is useful for
CC breeding plants with variable flower colours.
XX Sequence 1921 BP; 633 A; 296 C; 438 G; 554 T; 0 other;
SQ

```

## Alignment Scores:

```

Pred. No.: 4.11e-157 Length: 1921
Score: 1557.00 Matches: 305
Percent Similarity: 81.82% Conservatives: 46
Best Local Similarity: 71.10% Mismatches: 68
Query Match: 68.59% Indels: 10
DB: 19 Gaps: 2

```

US-09-847-081b-2 (1-440) x AAV16949 (1-1921)

```

Qy 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20
Db 406 GTTAACATGCTATTTGTACGCTATGGTGTGTTCCCGAGTTCTCAAGTTTTCAGTGGC 465
Qy 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 466 AATGTTTCTTGGAGCCAAATCGAGAAAGT-----TACCAT 501
Qy 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGlnArgTrp 60
Db 502 TTTTCGGATAAAAGTTTATGTACATGGAAGAGTTAAGAAAGTAGACACCAAGCGCT 561
Qy 61 AsnPheGlySerLeuIleAlaAspProArgTrpSerCysLeu-----GlyGlySerArg 78
Db 562 AGATCACGTTATGGGTTGGAGATTGTAGTTTCATTTGCTTGAGAGAGTCTGGATTAGAG 621
Qy 79 ThrGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlu 98
Db 622 ACCCGGGAAGAAGATTATCGTATCTCCAGTATTATAGTACCCCGGCGAGAGA 681
Qy 99 MetThrValSerSerGluLysLysValTrpAspValValLeuLysGlnAlaLeuVal 118
Db 682 ATGACGATGACATCAGACAAAGGTTTATGATGTCGTTTAAAGCAAGCAGCTTTGATT 741
Qy 119 LysArgGlnLeuArgSerThrAspLeuGluValLysProAspIleValProGly 138
Db 742 AATAGACAGTTGAGGCTAGAGAAATTTGGAGGTGAAACCGGACATTTATTTGCCAGGA 801
Qy 139 AsnLeuGlyLeuLeuSerGluAlaTrpAspArgCysGlyGlyValCysAlaGluTrpAla 158
Db 802 AACGGCAACGCTGTAATGAAGCTTATGATCGGTGCGAGAGATGTCGTAATATGCC 861
Qy 159 LysThrPheTrpLysLeuMetThrProGluArgArgArgAlaIleTrpAla 178
Db 862 AAGTCATTCCTACTGGGNAACGAGCTCATGACCCGAGGAGCGGTTTAGCTATCTGGCG 921
Qy 179 IleTrpValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIle 198

```

```
Db 922 ATATATCTATGGTGTAGGACGACAGTACGCTTGTGATGGGCTAACCGCTCAACATA 981
QY 199 ThrProGlnAlaLeuAspArgTrrpGluThrArgLeuGluAspIlePheSerGlyArgPro 218
Db 982 AATCCACCCGCTTAGATAGTGGGAAGCAAGATTAGAAGATCTTTCAAAGGCAACCT 1041
QY 219 PheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGln 238
Db 1042 TTTGATATGCTTGTGCTGCTTATCTGATACCATACCAAGATATCTGTGGACATCCAG 1101
QY 239 ProPheArgAspMetIleGluGlyMetArgMetAspLeuTrrpLysSerArgTrrpLysThr 258
Db 1102 CCATTTAGACATATAGTAGAAGCAATGCCGATGCTGAGAAATTCAGATACAGAAAT 1161
QY 259 PheAspGluLeuTrrpLysCysTrrpValAlaGlyThrValGlyLeuMetSerVal 278
Db 1162 TTCGATGAGCTGATCTTTACTGCTATTATGCTGGTACAGTTGGCTTGATGAGTGA 1221
QY 279 ProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAla 298
Db 1222 CCAGTATGGCATGTCACCTGAATCTAAGGCAACACAGAAAGTGTGTATATGTCAGCT 1281
QY 299 LeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAla 318
Db 1282 TTATCTTTGGGATCGCGAACACGACTGACTACATTTCTAAGGATGTTGGAGAAGTGA 1341
QY 319 ArgArgGlyArgValTrrpLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGlu 338
Db 1342 AGAAGAGGAAGAGTGTACCTACCTCAAGATGAATTAGCAACAGCAGGTTTATCAGATGAG 1401
QY 339 AspIlePheAlaGlyArgValThrAspLysTrrpArgAsnPheMetLysLysGlnIleGln 358
Db 1402 GACATTTTCTGGAAGATTCAGACAATGAGGATTTTATGAAGCAAGCAAAATCAA 1461
QY 359 ArgAlaArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSer 378
Db 1462 AGGCTAGAAATCTATGATGATGATGAGCAAAAAGTCCCGGACTCAGCTCCGCGAGC 1521
QY 379 ArgTrrpProValLeuThrAlaLeuLeuLeuTrrpArgLysIleLeuAspGluIleGluAla 398
Db 1522 AGATTGCTGTGTGGGCGAGCTTGTCTTTTATAGAAAATATTTGGATGATAGAAGCA 1581
QY 399 AsnAspTrrpAsnAsnPheThrArgArgAlaTrrpValSerLysProLysLysLeuLeuThr 418
Db 1582 AATGACTACAACTTCACAAAAGGGCTTATGTAAACAAGCAAGCAAGCTATTAGCT 1641
QY 419 LeuProIleAlaTrrpAlaLysSerLeu 427
Db 1642 ATGCTGTAGCATGTGCCAGTCTCTC 1668

RESULT 13
AAC48162
ID AAC48162 standard; DNA; 1566 BP.
AC AAC48162;
XX
XX
XX 18-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 56485.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
```

PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Alignment Scores:		
Pred. No.:	1,25e-155	Length: 1566
Score:	1542.00	Matches: 318
Percent Similarity:	81.86%	Conservative: 43
Best Local Similarity:	72.11%	Mismatches: 51
Query Match:	67.93%	Indels: 30
DB:	21	Gaps: 8
US-09-847-081B-2 (1-440) x AAC48162 (1-1566)		
Qy	1 MetSerMetSerValAlaLeuLeurpValVal-----SerProThrSerGluValSer 18	
Db	153 ATGTCTTCTTCGTAGCAGTGTTATGGTGTCTACTTCTTAATCCAGACCCAATG 212	
Qy	19 AsnGlyThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArg 38	
Db	213 AACAATTGGGTGGTGGTA-----AGGTTCTAGAATCTTCTAGA 251	
Qy	39 PheLeuAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGly-GlyArgGI 58	
Db	252 CTGTTCTCTCCTTGTGCAG-----AATCAGAGACTAAACAAGSTAAGAAGAAG 299	
Qy	58 nArgrTrpAsnPheGlySerLeulleAlaAspProargTyrrSerCysLeuGlyGlySerAr 78	
Db	300 CAGATACCACCACTTGGAGTTCTTCTT-----TTGTAAGGAACCGA 338	
Qy	78 gThrGLuLySgLYSerThrPheSerValGlnSerSerLeuValAlaIaSerProAlaGLYl 98	
Db	339 AGTAGAAGAAATT-GGTGTT-----GTGCTTCAAGCTAGTAGCAAGTCCTTCTCGAGA 391	
Qy	98 uMethThrValSerSerGluLySlysValTyrrAspValValLeuLySglnAlaAlaLeuVa 118	
Db	392 GATAGCTCTTTCATCTGAAGAGAGGTTTACAATGTTGTGTGAACAAGTCGCTTGTGT 451	
Qy	118 lLysArgGlnLeuArgSerThrAsp---AspLeuGluValLys-----ProAspIleVa 135	
Db	452 GAACAACACAGCTAAGGTCTTCTTCTTATCACCTTGATGTGAAGAAACCAACAGATCTGT 511	

QY 135 lvalProGlyAsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAl 155  
DB 512 TCCTCTGGAGTTGAGTTTGGTGAAGCTTATGATCATCGGTGAAGTTGGCG 571  
QY 155 aGluTyrAlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAl 175  
DB 572 TGAATATGCTAAGACGCTTTTATCTTGGAACTTTGCTTATGACACCGAAGCGAAGGC 631  
QY 175 alletPAlaIleTyrValTyrCysArgArgThrAspGluLeuValAspGlyProAsnAl 195  
DB 632 GATTTGGGCAATCTAGTTTGGTGAAGAACTGATGAACCTGTGGATGGGCCAAATGC 691  
QY 195 aserHisIleThrProGluAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSe 215  
DB 692 TTCACATATACTCCCATGCTTTAGATAGATGGGAAGCAAGTTAGAAAGATCTTTCCG 751  
QY 215 rGlyArgProPheAspMetLeuAspAlaLeuSerAspThrValSerArgPheProVa 235  
DB 752 TGTCTGCTCTTTCGATATGCTTGTCTGCTGCTGATACAGTTGCTAGATACCCGCT 811  
QY 235 lAspIleGlnProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerAr 255  
DB 812 CGATATTCAGCCATTTCCGACATGATCGAAGGAATGAGATGGACTTGAAGAAATCGAG 871  
QY 255 gTyrLysThrPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLe 275  
DB 872 ATACCGAAGCTCGATGATCTATACCTTTACTGCTACTGCTGGAACCGCTCGGAT 931  
QY 275 uMetSerValProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTy 295  
DB 932 GATGAGCTTCCGGTTATGGGAATCATCTAGTCAAGCAACCAACCAAGTCTTTA 991  
QY 295 rAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGl 315  
DB 992 CAACGCTGCTTGGCTTGGTATGCCAATCACTAGCTTACTTAACATACTCAGAGACGTAGG 1051  
QY 315 YGluAspAlaArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLe 335  
DB 1052 CGAAGATCGCAGAGAGAGAGGTTTATCTGCTCAGGATGATTTGCTCAGCTCGCTGCT 1111  
QY 335 uSerAspGluAspIlePheAlaGlyArgValThrAspLysTyrPArgAsnPheMetLysLy 355  
DB 1112 TTCAGATGAGACATATTCGCCGGAAGTAAGTATGATAAATGAGAACTTCATGAAAT 1171  
QY 355 sGlnIleGlnArgAlaArgLysPheAspGluSerGluLysGlyValThrGluLeuAs 375  
DB 1172 GCAGCTTAACGAGCAAGATGTTCTTCGACGAAGCTGAGAAAGGCGTCACCGAGCTCAG 1231  
QY 375 pSerAlaSerArgTyrProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGl 395  
DB 1232 TGCCGCTAGCAGATGCGCTGATGGCTTCATTGCTTATTGACAGAGAACTACTGGACGA 1291  
QY 395 uIleGluAlaAsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLy 415  
DB 1292 GATTGAGCGCAATGATTACCAATTTTACTAAGAGAGCTTATGTTGGGAAGTCAAGAA 1351  
QY 415 sLeuLeuThrLeuProIleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSe 435  
DB 1352 AATTGAGCTTCCCATGCTTATGCTAATCAGTACTA-----AAGACTTCAAG 1402  
QY 435 r 435  
DB 1403 T 1403  
RESULT 14  
AAC35120  
ID AAC35120 standard; DNA; 1703 BP.  
XX  
AC AAC35120;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9067.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW Protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
PN EP1033405-A2.  
XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.



```
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1,41e-155
Score: 1542.00
Percent Similarity: 79.2%
Best Local Similarity: 70.38%
Query Match: 67.93%
DB: 21
Length: 1703
Matches: 316
Conservative: 40
Mismatch: 47
Indels: 46
Gaps: 8

US-09-847-081b-2 (1-440) x AAC35120 (1-1703)
QY 1 MetSerMetSerValAlaLeuLeuTrpValVal-----SerProThrSerGluValSer 18
Db 290 ATGTCTTCTCTGTAGCAGTGTATGGTTGCTACTTCTCTCTAAATCCAGACCCCAATG 349
QY 19 AsnGlyThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerArg 38
Db 350 AACAAATGTGGTGGTA-----AGGTTCTAGAAATCTTCTAGA 388
QY 39 PheLeuAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGln 58
Db 389 CTGTTCTCTCCTGTGCAGAA-----CAG 412
QY 59 ArgTrpAsnPheGlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArg 78
Db 413 AGACTAACAAAGGTAAGAGACAGATACCAACTTGGAGT----- 454
```

```
QY 79 ThrGluLysGlySerThrPhe-----SerValGlnSer 89
Db 455 -----TCTCTTTTGAATGACCAAGTAGAAGATTGGTGTGTGCTTCA 502
QY 90 SerLeuValAlaSerProAlaGlyGluMetThrValSerSerGluLysLysValThrAsp 109
Db 503 AGCTTAGTAGCAAGTCCCTCTCTGGAGAGATAGCTCTTTCATCTCAAGAGAAGGTTTACAAAT 562
QY 110 ValValLeuLysGlnAlaAlaLeuValLysArgGlnLeuArgSerThrAsp---AspLeu 128
Db 563 GTTGCTTGAACAGCTCTTGGTGAACAACAGCTAAGTCTCTCTCTATGATCT 622
QY 129 GluValLys-----ProAspLeuValProGlyAsnLeuGlyLeuSerGluAla 146
Db 623 GATGTGAAGAACCACCAAGATGTTGCTCTCTGGAGTTGAGTTGTTGGTGAAGCT 682
QY 147 TyrAspArgCysGlyValCysAlaGluTyrAlaLysThrPheTyrLeuGlyThrLys 166
Db 683 TATGATCGATGGGTGAAGTTGGCTGAATATGCTAAGACCTTTTATCTTGAACCTTG 742
QY 167 LeuMetThrProGluArgArgAlaIleTyrAlaIleTyrValTyrCysArgArgThr 186
Db 743 CTTATGACACCCGAAGGCGAAGGATTTGGCTATCTAGCTTTGGTGTAGAAGAACT 802
QY 187 AspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAlaLeuAspArgThr 206
Db 803 GATGAACCTTGTGGATGGCGCAATGCTTCAATATAACTCCCATGCTTTAGATAGATGG 862
QY 207 GluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspAlaLeu 226
Db 863 GAAGCAAGTTAGAGATCTTTCCTGCTGCTCTTTCGATATGCTTGCATGCTGCTCTC 922
QY 227 SerAspThrValSerArgPheProValAspIleGlnProPheArgAspMetIleGluGly 246
Db 923 GCTGATACAGTTGCTAGATACCTGCGATATTCAGCCATTTCCGACATGATCGAAGGA 982
QY 247 MetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGluLeuTyrLysCys 266
Db 983 ATGAGAAATGGACTTGAAGAAATCGAGATACCAAGCTTCGATGATCTATACCTTACTGC 1042
QY 267 TyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaProGlu 286
Db 1043 TACTAGTCTGCGAACCCTCGATGATGAGCGTTCGCGTTATGGATCGATCTCAAG 1102
QY 287 SerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGln 306
Db 1103 TCGAAAGCAACACCCGAAAGTGTTCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162
QY 307 LeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTyrLeuPro 326
Db 1163 CTTACTAATACATCTAGACAGCTGAGCGAAGATGCGAAGAGAGGAGGTTTATCTGCT 1222
QY 327 GlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArgValThr 346
Db 1223 CAGGATGAATGGCTCAGCGTGGTCTCTCAGATGAAGACATATTCGCCGGAAGATTA 1282
QY 347 AspLysThrPArgAsnPheMetLysLysGlnIleGlnArgAlaArgLysPheAspGlu 366
Db 1283 GATAAATGAGAACTTCATGAAATGCAGCTTAACAGCAAGAAATGTTCTTCGACGAA 1342
QY 367 SerGluLysGlyValThrGluLeuAspSerAlaSerArgTyrProValLeuThrAlaLeu 386
Db 1343 GCTGAGAAGGCGCTACCGAGCTCAGTCCGCTAGCAGATGCGCTGTATGGCTTCATTG 1402
QY 387 LeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsnAsnPheThrArg 406
Db 1403 CTATTGTACAGGAGATATGACGAGATTTGAAGCGAATGATTAACCAATTTTACTAAG 1462
QY 407 ArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAlaTyrAlaLysSer 426
Db 1463 AGAGCTTATGTGGGAAAGTCAAGAAATTCAGAGCTTTGCCATTCGCTTATGCTTAATCA 1522
```

```
QY 427 LeuValProAsnArgThrSerSer 435
Db 1523 GFACTA-----AAGACTTCAAGT 1540
RESULT 15
AAV16948
ID AAV16948 standard; cDNA to mRNA; 2085 BP.
XX
AC AAV16948;
XX
DT 06-JUL-1998 (first entry)
XX
DE Nucleic acid encoding phytoene synthase 1.
XX
KW Phytoene synthase; breeding; variable flower colour; ds.
XX
OS Gentiana lutea.
XX
FH Key Location/Qualifiers
FT CDS 586..1866
FT /tag= a
XX
PN JP10084966-A.
XX
PD 07-APR-1998.
XX
PF 17-SEP-1996; 96JP-0245107.
XX
PR 17-SEP-1996; 96JP-0245107.
XX
PA (INAT-) IWATE KEN.
XX
DR WPI; 1998-264853/24.
DR P-PSDB; AAW46961.
XX
PT phytoene synthase gene - useful for breeding plant of variable.
XX flower colour
XX
PS Claim 1; Pages 5-7; 15pp; Japanese.
XX
CC The present sequence encodes phytoene synthase 1. It was isolated from
CC a cDNA library prepared from mRNA extracted from the petals of Gentiana
CC lutea. The nucleic acid sequence was amplified from the library using
CC PCR primers AAV16952-53. The phytoene synthase gene is useful for
CC breeding plants with variable flower colours.
XX
SQ Sequence 2085 BP; 654 A; 320 C; 479 G; 632 T; 0 other;
```

## Alignment Scores:

Pred. No.:	2.89e-154	Length:	2085
Score:	1531.00	Matches:	310
Percent Similarity:	82.56%	Conservative:	45
Best Local Similarity:	72.09%	Mismatches:	64
Query Match:	67.44%	Indels:	12
DB:	19	Gaps:	3

US-09-847-081B-2 (1-440) x AAV16948 (1-2085)

```
QY 1 MetSerMetSerValAlaLeuLeuTyrValSerProThrSerGluValSerAsnGly 20
Db 580 GTTACATGCTCATTTGTGTAGCGTATGGGTGTTTCCCGAGTTCTCGAAGTTTGTAGTGGC 639
QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 640 AATGTTTTCTTGGAGCCATTTCGAGAACT-----TACCAT 675
QY 41 AlaArgAspArgAsnLeuMetThrPasnGlyArgIleLysLysGlyGlyArgGlnArgTrp 60
Db 676 TTTTCGGATAAAGTTTAATGTACAAATGGAAGAGTTAAGAAAATAGACACCAAGCGCT 735
QY 61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
Db 736 AGATCACGTTATGGGGTTGGAGATTTCAGTTTCATTTTGTCTTGAGAGAGTCTGATTAGAG 795
```

Search completed: April 5, 2003, 03:09:28  
Job time : 306 secs

```
QY 81 ---LysGlySerThrPheSer-ValGlnSerSerLeuValAlaSerProAlaGlyGluMe 99
Db 796 ACCCGGGGAAGAAGATTATCGGTATCTCCATATTAGCTACCCCGCAGGAGAAAT 855
QY 99 tThrValSerSerGluLysLysValTyrAspValValLeuLysGlnAlaAlaLeuVally 119
Db 856 GACATG-ACATCAGAGCAAAAGGTTTATGATGTCGTTTTAAGCAACGACGCTTGATTA 914
QY 119 sArgGlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAs 139
Db 915 TAGACAGTTGAGTCTAGAGAAATTTGGAGGTGAACCCGACATTTATTTGCCAGGAAA 974
QY 139 nLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLy 159
Db 975 CGCAACGCTGTTGAATGAAGCTTATGATCGGTGTCGAGAAAGTATGTCTGAATATGCCAA 1034
QY 159 sThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAla-- 178
Db 1035 GTCATCTCTAGTGGGAACCCAGCTCATGACCGGAGAGCGGTTTAGCTATCTGGGCCGA 1094
QY 179 -IleTyrValTrpCysArgThrAspGluLeuValAspGlyProAsnAlaSerHisIi 198
Db 1095 TATATATGTTATGGGTAGGAGACAGATGAGCTTGTGTGGCCTAACCGCTCACACAT 1154
QY 198 eThrProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgPr 218
Db 1155 AAATCCACCCGCGTTAGTAGTGGGAAGCAAGATCAGAAGTGTTTTCACAGGGCAACC 1214
QY 218 oPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGl 238
Db 1215 TCTGGATATGCTGATGCTGCTTATCTGATACCATTTACCAAGTATCTCTGGACATCCA 1274
QY 238 nProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysTh 258
Db 1275 GCCATTTAGAGATATGATAGAGGATGCGGATGGATCTGAAGNAATCGAGATACAAGAA 1334
QY 258 rPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerVa 278
Db 1335 TTTGATGAGCTGATCTTTACTGCTATTATGTGGCTGGCACAGTTGGCTTGATGAGTGT 1394
QY 278 lProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAl 298
Db 1395 ACCAGTAATGGGATTCACCTGAATCTAAGGCAACACAGAAAGTGTGTATGATGCAGC 1454
QY 298 aLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAl 318
Db 1455 TTTATCTTTGGGATCCGAGACCAGCTGACTACATTTCTAAGGGATGTTGGAGAAATGC 1514
QY 318 aArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGl 338
Db 1515 AAGAAGAGGGAGAGTGTACCTACTCAAGATGAATTAGCACAAGCAGGTTTATCAGATGA 1574
QY 338 uAspIlePheAlaGlyArgValThrAspLysTrpArgAsnPhemMetLysGlnIleGl 358
Db 1575 GGACATTTTTCGGAAGAAAGTTACAGACAAATGGAGGATTTTATGAAGAAGCAAAATCAA 1634
QY 358 nArgAlaArgLysPheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSe 378
Db 1635 AAGGGCTAGAAATTTCTATGATGATGCAGAGAAAGGAGTCCCGAAGTCTCAGCTCCGCGAG 1694
QY 378 rArgTrpProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAl 398
Db 1695 CAGATTGCTGTGTGGCAGCGTTGCTTTTATAGAAAAATATTTGGATCAGATAGAGC 1754
QY 398 aAsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuTh 418
Db 1755 AAATGACTACAAATTTCCCGCAGAGGGCTTATGTAAACGAGGCGGAAGGCTATTAGC 1814
QY 418 rLeuProIleAlaTyrAlaLysSerLeu 427
Db 1815 TATGCCCTGTAGCATGTGCCAAGTCTCTC 1842
```